ECMWF Feature article

from Newsletter Number 114 – Winter 2007/08

METEOROLOGY

Probability forecasts for water levels at the coast of The Netherlands

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/gpsn56s02c

This article appeared in the Meteorology section of ECMWF Newsletter No. 114 – Winter 2007/08, pp. 23-29.

Probability forecasts for water levels at the coast of The Netherlands

Hans de Vries, KNMI, De Bilt, The Netherlands

Accurate storm surge forecasts are of paramount importance for a country which, like The Netherlands, has an important part of its economic activity and population concentration close to the coast in a part of the country which is near or even below sea level. The primary usage is for coastal defence and the operation of the storm surge barriers in the Oosterschelde and the Nieuwe Waterweg, the entrance to the harbour of Rotterdam. But the forecasts are also used for assisting shipping traffic in and out of the Rotterdam harbour and by inland water management boards to assess their ability to sluice water into the sea at low tide in case of heavy rainfall.

Forecasts are issued by KNMI, the Royal Netherlands Meteorological Institute, in close cooperation with the Storm Surge Warning Service (SVSD) of Rijkswaterstaat, the Public Works and Water Management Authority. When there is a severe storm surge, the SVSD takes the final responsibility for the water level forecasts that are issued and warns other appropriate authorities.

The standard tool available for numerical operational storm surge forecasts is the two-dimensional shallow-water model WAQUA/DCSM98. This model covers all of the NW European Continental Shelf, which includes the North Sea, with a resolution of 8 km. Figure 1 shows the area covered by the model and the water depths. The atmospheric forcing for day-to-day forecasts for up to 48 hours ahead is provided by the 10-metre wind and mean sea level pressure from KNMI's operational HIRLAM limited area model. This system provides sea level forecasts four times per day. A steady-state Kalman filter is used to assimilate water level observations from tide gauges on the east coast of Britain and the Dutch coast.

Recently, there has been growing attention to the increased possibility of and the vulnerability to flooding due to the effects on sea level and storm climate caused by global warming. This has prompted the extension of the forecast range, together with the need for information on the uncertainty of the forecasts. ECMWF's Ensemble Prediction System (EPS) is an obvious candidate to provide input for such forecasts. A comparatively simple model which calculates surges using a statistical technique had already been coupled to the EPS to produce forecast plumes. But now, the increased capacity of computers allows for a water level ensemble based on the numerical two-dimensional storm surge model, which is generally more accurate and also more flexible in its output. The results of this ensemble are being used to produce probability forecasts for water levels in selected locations.

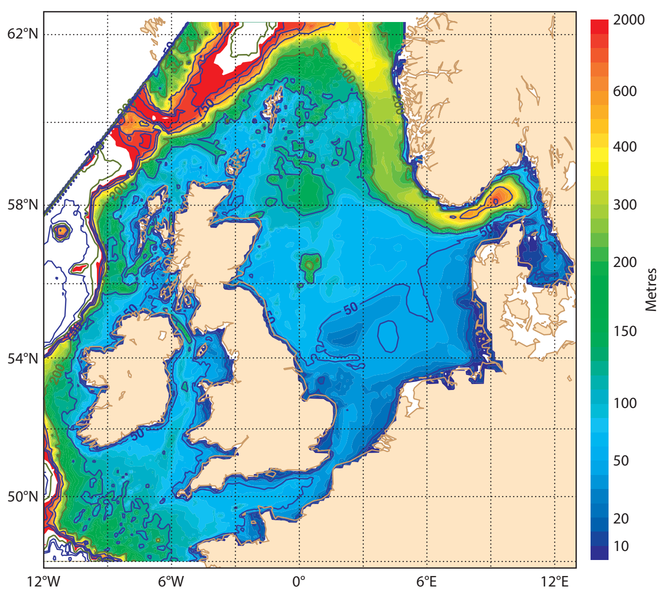


Figure 1 Depth map of the WAQUA/DCSM98 storm surge model.

doi:10.21957/gpsn56s02c

The forecasting system

WAQUA/DCSM98 with input from the EPS has been running in real-time experimental mode since February 2007. With data from the winter 2003/04, hindcasts have been run to establish the calibrations which are needed to generate probability forecasts from the ensembles. Results for eight locations along the coast of The Netherlands are presented on the Internet. Figure 2 gives an example of a forecast.

Probabilities derived from the ensembles are represented by bars with varying thickness and colours to indicate the range. Astronomical extremes are given as black crosses and observations, if already available, as red circles. Also shown are the levels defined by the SVSD to indicate the severity of the surge. The probabilities never span the full 100% range. One reason is that the derivation of the probabilities explicitly allows for a small fraction of the forecasts to lie outside the ensemble. Also ensemble members might be so close together, especially at the edge of the ensemble, that parts of the probability density function collapse. For forecasts less than 48 hours ahead, lack of spread in the ensemble makes more of the probability fall outside the ensemble and generally not more than the 25% to 75% interval can be determined.

The forecasts are being used, for example, for medium-range planning of the personnel required in case of a severe storm. But there is also a keen interest from inland water management boards. They use the information on predicted low-tide levels for medium-range planning of the regulation of inland water levels when heavy rainfall is forecast.

In November 2007 a storm occurred which reached the alarm level for the whole of the Dutch coast, except for Vlissingen in the south where it fell just a few centimetres short. In Hoek van Holland the water reached the highest level since the disaster in 1953. It was the first time in more than 30 years that warnings and alarms had to be issued for the entire Dutch coast. The storm surge barriers in the Oosterschelde and the Maeslantkering and Hartelkering near Rotterdam had to be closed, the latter two for the first time in the event of a storm surge. The high levels were successfully captured as far as five days ahead. Figure 2 shows the forecast for Hoek van Holland three days ahead. Such forecasts are used to decide whether to close the barriers near Rotterdam.

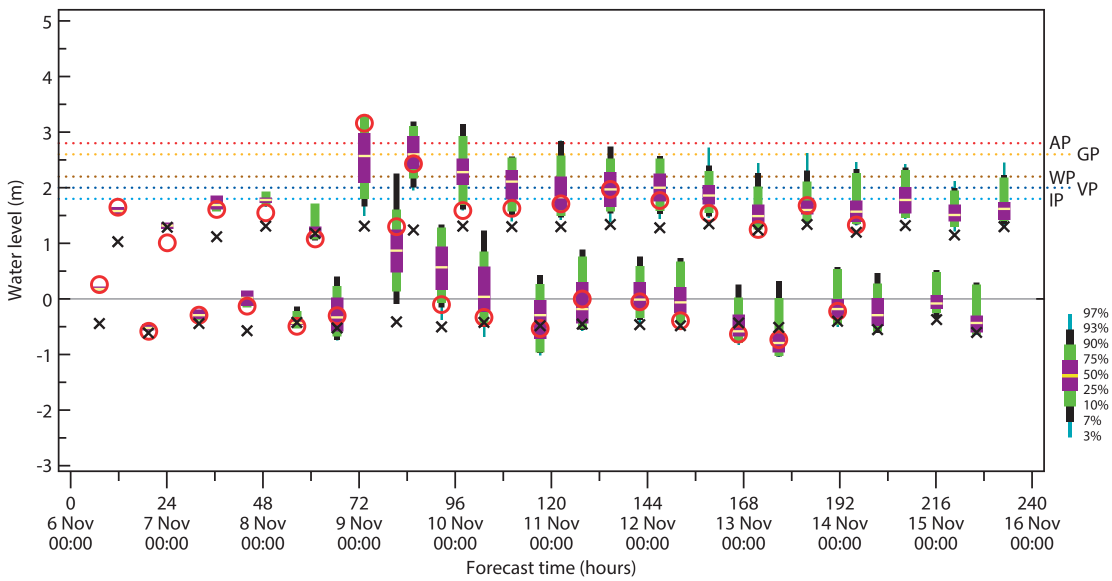


Figure 2 Example of a probability forecast for Hoek van Holland (00 UTC on 6 November 2007). Astronomical extremes are given as black crosses and observations as red circles. The horizontal lines indicate the levels defined by the SVSD to indicate the severity of the surge: information (IP), pre-warning (VP), warning (WP), storm surge (GP) and alarm level (AP).

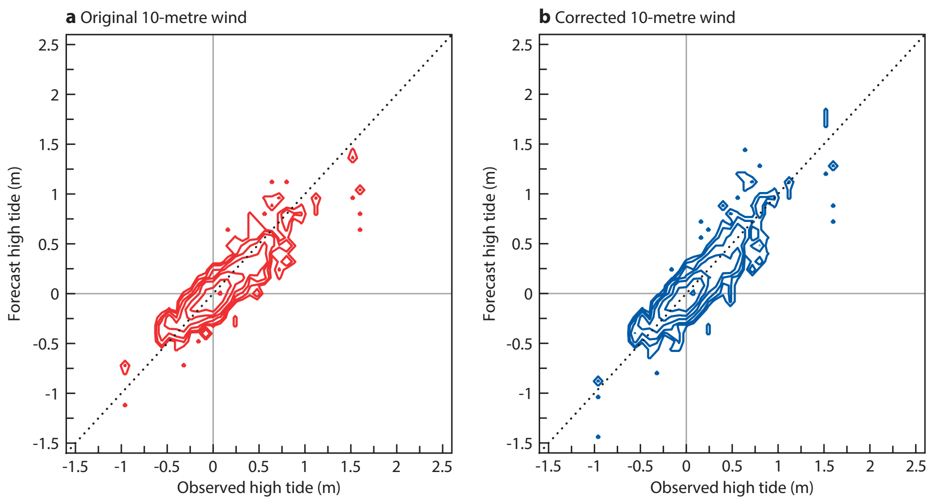
Calibration of deterministic forecasts

The transformation of the 'raw' ensemble forecasts into probability forecasts and an evaluation of their quality involved several steps.

The first step was the validation of deterministic water level forecasts with input from the deterministic high-resolution ECMWF atmospheric model. This was done for the period September 2003 to April 2004, a recent winter which contained a reasonable amount of storm surges. As one of the validation tools, the correlation between forecast tidal extremes and observations is used. Figure 3(a) shows the correlation for IJmuiden. This is essentially a scatter plot which has been transformed into a two-dimensional histogram. The contour lines indicate the density of the points, starting from 10% below the highest level and go logarithmically down to single points. The deviation from the main diagonal, as seen in Figure 3(a), indicates a systematic under-prediction of surges.

As storm surge forecasts with input from HIRLAM do not suffer from such systematic under-prediction, a comparison was made between the 10-metre wind fields (i.e. the main input to the storm surge model) from HIRLAM and the ECMWF model. This comparison is shown in Figure 4 for a collection of wind fields for the same forecast time from stormy periods. As in Figure 3, this gives the densities of the two-dimensional histogram of the wind speeds over the North Sea from HIRLAM and the ECMWF model interpolated to the grid points of the storm surge model. It shows an under-prediction for higher wind speeds. To quantify this, averages and standard deviations of the densities have been calculated perpendicular to the main diagonal; these are shown by the dashed (average) and red lines (average ± standard deviation). For wind speeds below 8 $m s^{-1}$ the averages lie close to the main diagonal, but the excess above this is on average 14% less in the ECMWF model than in HIRLAM. The green line in Figure 4 summarises this. Similar plots for wind directions show no systematic differences.

A rerun of the storm surge model shows that applying a wind correction of +14% on the excess of 8 $m s^{-1}$ indeed removes the systematic under-prediction of the results. This is shown in Figure 3(b). For the control run of the EPS the same correction is found.



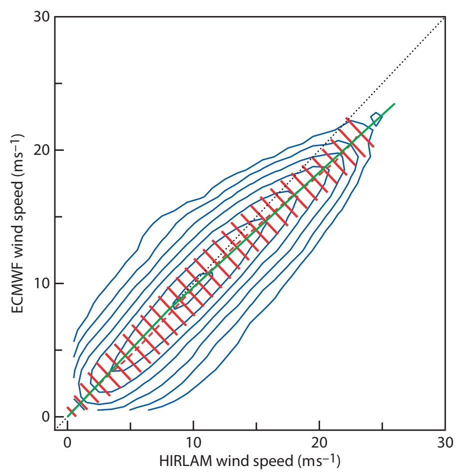


Figure 3 (Above) Validation of deterministic forecasts from T+48 to T+96 for high tides for IJmuiden with (a) original 10-metre wind and (b) corrected 10-metre wind.

Figure 4 (Left) Correlation between HIRLAM and ECMWF winds for stormy periods. Averages and standard deviations of the densities calculated perpendicular to the main diagonal are shown by the dashed (average) and full red lines (average ± standard deviation).

 doi:10.21957/gpsn56s02c

Calibration of the ensemble

To assess the quality of the sea level ensemble and derive probabilities, forecast ensembles have been regenerated for the winter 2003/04, both with and without the correction of the wind speed described earlier. The ensembles are studied by means of Rank Histograms, also known as Talagrand Diagrams. To construct these, a ranking from low to high is established for the forecasts for each of the tidal extremes for each of the sea level ensembles. This gives unequally spaced forecast intervals for each of the tides, which are unique for every ensemble forecast and for every tidal extreme.

For an ideal forecasting system the realization of the sea level is equally likely to fall in either of these intervals, including two for everything below or above the ensemble. The Talagrand Diagram is a histogram of these realizations. Its shape characterises the forecasting system. In the ideal case where all intervals are populated equally, the histogram is flat. An ensemble system with too little variation between the members leads to realizations which are often at the edges or outside of the ensemble, and hence gives a u-shaped histogram. However, too much variation gives under-population of the edges and hence a bell shape. A skew diagram indicates under- or over-prediction. The Talagrand Diagram can be summed to get a so-called Cumulative Rank Histogram (CRH). This gives the relation between ensemble probabilities and observed frequencies. An additional advantage of the CRH over the Talagrand Diagram is that the former smooths out the scatter between the different intervals and hence gives a clearer characterisation of the system.

Figure 5 gives an example of a Talagrand Diagram and the corresponding CRH. The diagrams have been produced from all high tide forecasts from 60 to 96 hours ahead for Hoek van Holland. The red curves are without the correction of the 10-metre wind and the black curves include this correction.

The under-population of the lower and over-population of the higher forecast intervals indicate that the ensemble considerably underestimates the surges. Although the correction of the 10-metre wind works the right way and brings down the fraction of the forecasts where the observation is higher than the ensemble, from 12% to 8% in this case, it is not enough by far. Moreover, the under-population of the lower forecast intervals does not change significantly. An explanation for that can be found in the observation that the lower surges that populate these bins are caused by lower winds for which the correction is smaller.

These features are typical for the whole range of the Dutch coast and all forecast ranges from 48 hours ahead. For ranges up to 48 hours the diagrams are characterised by too little variation in the ensemble members. This, however, is a well-known feature of the EPS. In order to overcome the under-prediction, various correlations between ensemble parameters and observations have been investigated. The most promising is the relation between the standard deviation of the ensemble and the difference of the ensemble mean from the observation as illustrated in Figure 6; this figure shows an increase of the average difference with ensemble width. For positive surges there is an increasingly negative difference between the ensemble mean and the observations and for negative surges this difference gets more and more positive. This trend points to a bifurcation in the probability density function for the larger surges, positive or negative. A number of ensemble members does catch this surge, but others do not and the standard deviation of the ensemble as a whole increases.

From this relation a simple correction has been constructed to bring the average difference back to zero by shifting the ensembles according to their standard deviation. The Talagrand Diagrams, regenerated with this correction, exhibit a very much improved shape and the systematic under-prediction has been removed, as shown in Figure 7.

Skill of the forecasts

The corrected CRHs can be used to calculate calibrated probabilities from the ensemble forecasts. To validate these, the Brier Skill Scores for the exceedance of different levels have been determined. The sample climatology has been used as reference. As an example, Figure 8 gives the Brier Skill Scores for the exceedance of a range of surge levels for 36 hours starting from day 2 and day 5 for IJmuiden for low (black curves) and high tides (red) and both with the corrected (solid) and the uncorrected (dashed) CRH diagrams.

For surges between -0.5 m and +0.8 m there is skill, even after day 5, when the corrected CRH is used. Outside this surge range there were only a few cases during the period, which makes the results statistically insignificant and the Brier Scores vary wildly.

For the results shown in Figure 8 the 10-metre wind correction was applied, but without the wind correction the results are similar. This means that, although the correction improves the deterministic forecast, it can be omitted for the generation of the probability forecasts. Therefore, it is not applied in the real-time ensemble system.

Because the 10-metre wind correction can be omitted for the generation of the probability forecasts, recalibration becomes easier. The ensembles do not have to be rerun once a new 10-metre wind correction has been established, but a new calibration for the ensemble can simply be derived from the forecasts that have already been generated.

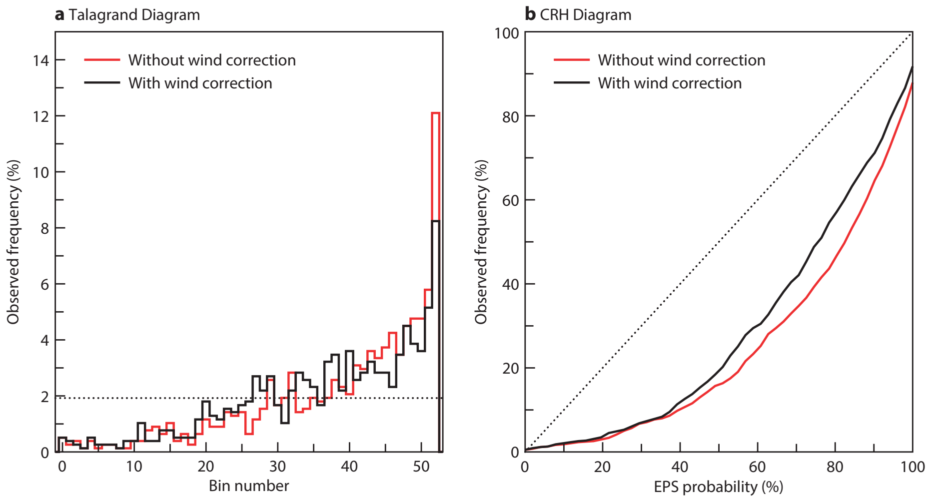


Figure 5 (a) Talagrand Diagram and (b) CRH diagram for all high tide forecasts from T+60 to T+96 for Hoek van Holland without corrections to the ensemble. The red lines are without the 10-metre wind correction, the black lines include the correction.

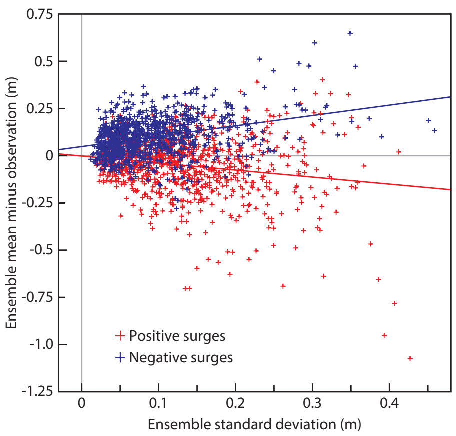


Figure 6 Ensemble mean minus observation versus ensemble standard deviation. Red crosses are from positive surges, blue crosses from negative surges.

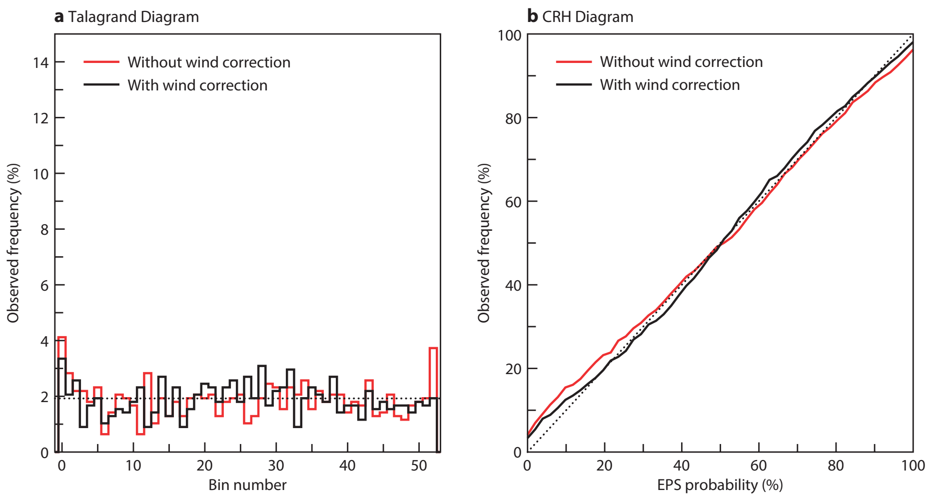


Figure 7 (a) Talagrand Diagram and (b) CRH diagram for all high tide forecasts from T+60 to T+96 for Hoek van Holland with corrections to the ensemble. The red lines are without the 10-metre wind correction, the black lines include the correction.

doi:10.21957/gpsn56s02c

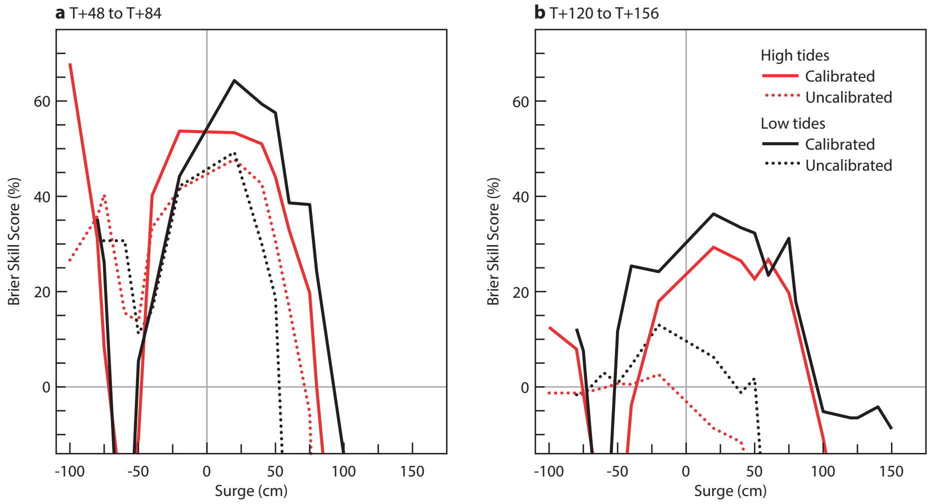


Figure 8 Brier Skill Scores for different forecast intervals for IJmuiden: (a) T+48 to T+84 and (b) T+120 to T+156. Black curves are for low tides, red curves for high tides. The solid curves are for the calibrated, dashed curves for uncalibrated probabilities.

Recalibration

A drawback of a calibrated forecasting system is that the calibration should be maintained, especially after changes in the underlying forecasting models. A major change of the deterministic atmospheric model and EPS occurred in February 2006 when the resolution of both was upgraded. The calibrations of both the deterministic water level forecasts and the water level ensemble were re-established from the data for September 2006 to March 2007. A fortunate advantage for this purpose was that this winter, like the one of 2003/04, contained several significant storm surges.

Corrections for the 10-metre wind from the deterministic model have reduced from 14% extra for the excess of 8 $m s^{-1}$ to 11% for the excess of 10 $m s^{-1}$. Also the calibrations for the ensembles have changed, but the under-prediction is still present and can be corrected for in the same way as described earlier.

Future developments

The medium-range probability forecasts have found an important application in providing guidelines for the planning of activities. For short-range forecasts there is also a desire for uncertainty information, which will help decision making, e.g. in case of a storm surge for the closing of storm surge barriers. The EPS unfortunately does not help on these time scales, but other possibilities include the use of an ensemble of storm surge models which are in use by various institutions around the North Sea or input for the storm surge model from a HIRLAM ensemble which is under development.

Further Reading

de Vries, H., 2007: Real-time water level probability forecasts. http://www.knmi.nl/~jwdv/WAQUA/EPS.

Gerritsen, H., H. de Vries & **M. Philippart,** 1995: The Dutch Continental Shelf Model. In *Quantitative Skill Assessment for Coastal Ocean Models,* Eds. D. Lynch and A. Davies, Vol. 47, Coastal and Estuarine Studies, American Geophysical Union.

Kok, C.J., 2001: Calibration of EPS derived probabilities. *Scientific Reports WR-2001-04,* KNMI, ISBN 90-369-2201-1.

Wilks, D.S., 2006: *Statistical Methods in the Atmospheric Sciences.* Vol. 91, International Geophysics Series, Academic Press, London, second edition, ISBN 13:978-0-12751966-1.